A Joint Location-Inventory Model

Zuo-Jun Max Shen • Collette Coullard • Mark S. Daskin
Department of Industrial & Systems Engineering, University of Florida, Gainesville, Florida 32611
Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois
Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois
shen@ise.ufl.edu • coullard@northwestern.edu • m-daskin@northwestern.edu

We consider a joint location-inventory problem involving a single supplier and multiple retailers. Associated with each retailer is some variable demand. Due to this variability, some amount of safety stock must be maintained to achieve suitable service levels. However, risk-pooling benefits may be achieved by allowing some retailers to serve as distribution centers (and therefore inventory storage locations) for other retailers. The problem is to determine which retailers should serve as distribution centers and how to allocate the other retailers to the distribution centers. We formulate this problem as a nonlinear integer-programming model. We then restructure this model into a set-covering integer-programming model. The pricing problem that must be solved as part of the column generation algorithm for the set-covering model involves a nonlinear term in the retailer-distribution-center allocation terms. We show that this pricing problem can (theoretically) be solved efficiently, in general, and we show how to solve it practically in two important cases. We present computational results on several instances of sizes ranging from 33 to 150 retailers. In all cases, the lower bound from the linear-programming relaxation to the set-covering model gives the optimal solution.

1. Introduction

We consider the design of a distribution system in which a single supplier ships product to a set of retailers, each with uncertain demand. To maintain appropriate service levels, safety stock is maintained at the retailers. To achieve risk-pooling benefits, inventory cost reductions, and possibly line-haul shipping benefits, some retailers may be chosen to serve as distribution centers. As a distribution center, a retailer receives shipments from the supplier and distributes directly to a number of other retailers. The safety stock for all retailers served by the distribution center is maintained at the distribution center. Therefore, less total safety stock is required than in the case in which every retailer maintains its own safety stock.

The problem is the following. Given a collection of retailers, each with uncertain product demand, determine how many distribution centers to locate, where to locate them, which retailers to assign to each distribution center, how often to reorder at the distribution center, and what level of safety stock to maintain to minimize total location, shipment, and inventory costs, while ensuring a specified level of service.

We assume that location costs are incurred when distribution centers are established. Line-haul transportation costs are incurred for shipments from the single supplier to the distribution centers. Local transportation costs are incurred in moving the goods from the distribution centers to the retailers. Inventory costs are incurred at each distribution center and consist of the carrying cost for the average inventory used over a period of time, as well as safety stock inventory that is carried to protect against stockouts that might result from the uncertain retailer demand. Inventory costs are also incurred at each retailer, but are not modeled explicitly for reasons outlined below.

The model we present was motivated by work that the second and third authors did for a local blood bank. The blood bank served approximately 30 hospitals in a multicounty region of suburban Chicago and

Vol. 37, No. 1, February 2003 pp. 40–55
0041-1655/03/3701/0040$05.00
1526-5447 electronic ISSN

northern Indiana. We were primarily concerned with the production and distribution of platelets, the most perishable and most expensive of all blood products. Platelets must be agitated frequently, once they are separated from the whole blood, making the inventory costs high. Also, platelets must be destroyed five days after they are harvested from the donor if they have not been transfused into a patient. By way of comparison, most other blood products last a month or more and require less careful and less expensive handling. Platelets are needed only in special cases. However, when they are needed, it is often the case that multiple units must be transfused at one time. Thus, the demand for platelets is highly variable.

At the time that we worked with this blood bank, each hospital maintained its own inventory of platelets. At many hospitals, particularly the largest users of platelets, hospital inventory managers were concerned primarily with not running out of platelets. Not surprisingly, at these hospitals many units of platelets outdated and had to be destroyed. Other hospitals tended to keep little inventory on hand and placed most orders on a STAT or emergency basis, often requiring the blood bank to send a special shipment of platelets to the hospital, thereby incurring large transport costs. The inventory management policies at some hospitals were so rudimentary that some hospitals both placed frequent STAT orders and often carried large quantities of platelets to the point at which they outdated. It was clear that a better system was needed.

Our idea was to establish regional centers at which platelets would be stored for a collection of nearby hospitals. By storing platelets at regional centers (located at a subset of the hospitals) and distributing platelets to nearby hospitals on an as-needed or daily basis, three objectives were likely to be achievable. First, and most importantly, we could use risk-pooling principles to reduce the necessary safety stock needed to protect against shortages. Second, the cost of emergency shipments could be reduced because platelets would be stored closer to each of the hospitals. Finally, the training for inventory managers in an improved system would be simpler and more cost effective because fewer individuals would be involved as the inventory would be maintained at a small number of regional distribution centers instead of being maintained at each individual hospital.

We present two different mathematical-programming models for this problem. The first is a location allocation model and the second is a set-covering model. To incorporate the shipment, inventory and penalty costs, our location allocation model includes nonlinear terms in the objective function. This nonlinearity appears in the pricing problem associated with the set-covering model. We show how to deal with this nonlinearity practically in two important special cases, which leads to a computational procedure that provides very nearly optimal solutions on a variety of instances of sizes up to 150 retailers. We also show how, in general, the pricing problem can be reduced to the problem of minimizing a submodular set function, which is known to be polynomially solvable, via the ellipsoid method.

The paper is organized as follows. Section 2 discusses relevant results in the literature. Section 3 presents our location allocation risk-pooling model. Section 4 presents our set-covering model and the results on its associated pricing problem. Section 5 reports computational results. Section 6 concludes with directions for future research.

2. Background

The existing models in the literature developed to aid in the design of the types of distribution systems we consider fall under inventory theory and location theory. The inventory theory literature tends to focus on developing and evaluating policies for supplying the distribution centers and policies for filling retailer orders; these policies are evaluated based on the resulting service levels (percentage of retailer orders that are filled within the acceptable waiting period), shipping costs, inventory costs, and shortage costs (costs incurred when an order cannot be filled within the acceptable waiting period). (See, for example, Hopp and Spearman 1996, Nahmias 1997, and Zipkin 1997.) This work tends to incorporate demand uncertainty. On the other hand, the location theory literature tends to focus on developing models for determining the number of distribution centers and their locations, as well as the distribution-center-retailer assignments. These decisions are evaluated based on resulting operational shipping costs and strategic location costs. With some notable exceptions, this work tends to ignore demand uncertainty. Daskin and Owen (1999) provide an overview of facility location modeling, as do the recent texts by Daskin (1995) and Drezner (1995). See Owen and Daskin (1998) for a review of recent dynamic and stochastic facility location problems.

Thus, the inventory literature tends to ignore the strategic location decision and its associated costs, whereas the location literature tends to ignore the operational inventory and shortage costs, as well as the demand uncertainty and the effects that reorder policies have on these and shipping costs.

A result due to Eppen (1979) is essential in our present work, namely that significant inventory-cost savings are achieved by grouping retailers, and thus capitalizing on the so-called "risk pooling effects." In particular, Eppen considers a single-period model, with N retail outlets. In Eppen's model, each retailer faces a normally distributed customer demand. Eppen compares the following two modes of operation for retailer supply management. In the *decentralized mode*, each retailer orders independently the quantity that minimizes that retailer's expected one-period holding and penalty costs. In the *centralized policy*, a single quantity is ordered (and presumably stored at a single distribution center), to minimize total expected one-period holding and penalty costs. Notice that the only costs considered in both modes are holding and penalty costs. In particular, costs to ship from the supplier to the distribution center and from the distribution center to the retailers in the centralized mode, and costs to ship from the supplier to the retailers in the decentralized mode, are ignored.

Eppen's simple model dramatically illustrates the available savings due to risk pooling. Assume customer demands are normally distributed with a mean μ_i and a standard deviation σ_i for customer i, and assume the correlation coefficient of demands at retailer i and j is ρ_{ij}. Eppen showed that the expected total cost under the decentralized mode is $K\sum_{i=1}^{N}\sigma_i$, whereas the expected total system cost under the centralized mode is

$$K\sqrt{\sum_{i=1}^{N}\sigma_i^2+2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\sigma_i\sigma_j\rho_{ij}},$$

where K is a constant depending on the holding and penalty costs and the standard normal loss function.

Thus, if the demands of the N retailers are independent, the optimal cost can be expressed by $K\sqrt{\sum_{i=1}^{N}\sigma_i^2}$, which is less than $K\sum_{i=1}^{N}\sigma_i$.

Lateral stock transshipment (Krishnan and Rao 1965), which involves monitoring the movement of stock between locations at the same echelon level of the supply chain, has been studied as a means of emergency or alternative supply after demand has been realized. When examining this problem researchers have traditionally made several assumptions regarding the cost structure; whereas these assumptions are reasonable in many situations, there are many others in which these assumptions do not hold (Herer and Rashit 1999). For a review of this problem, please refer to Herer et al. (1999).

3. The Location Allocation Risk-Pooling Model

Before formally formulating the model, we outline the inventory policy under which the system operates. As indicated above, we assume that the non-DC retailers maintain only a minimal amount of inventory. Therefore, we ignore this inventory in the model below. The retailers that are selected to operate as distribution centers (DCs) order inventory from the plant using an economic order quantity model (EOQ), which is an approximation to the (Q, r) model with *Type I Service* constraint (Hopp and Spearman 1996, Nahmias 1997). Axsater (1996) points out that for (Q, r) model,

> although it is, in general, relatively easy to derive the optimal solution, it is common in practice to use an approximate solution that is obtained in two steps. First the stochastic demand is replaced by its mean and the order quantity Q is determined by the deterministic EOQ formula. Given this batch quantity the optimal reorder point is determined in the second step (p. 830).

It is well known that this procedure provides a good heuristic solution. Zheng (1992) has shown that the maximum relative error is bounded by 0.125. Axsater (1996) shows a slightly better bound of 0.118. Numerical experiments of practical situations show that the performance of this heuristic is usually much better than the worst-case bound.

The frequency of orders and the order quantity at each DC is determined by the mean demand served by the DC which, in turn, is a function of the assignment of customers to the DC. In addition to this working inventory, each DC maintains a safety stock designed to protect against the possibility of stockouts during a lead time. With this background, the problem we consider is the following: Given a set I of retailers, each facing independent uncertain demand, decide how many distribution centers to locate, where to locate them, which retailers to assign to each distribution center, how often to reorder at the distribution centers, and what level of safety stock to maintain to minimize the total location, shipment, working inventory, and safety stock inventory costs.

To model this problem, we define the following notation:

Inputs and Parameters

- μ_i — mean (yearly) demand at retailer i, for each $i \in I$;
- σ_i^2 — variance of (daily) demand at retailer i, for each $i \in I$;
- f_j — fixed (annual) cost of locating a regional distribution center at retailer j, for each $j \in I$;
- $v_j(x)$ — cost to ship x units from the main supplier (the plant) to a regional distribution center located at retailer j, for each $j \in I$;
- d_{ij} — cost per unit to ship from retailer j to retailer i, for each $i \in I$ and $j \in I$;
- α — desired percentage of retailers orders satisfied;
- β — weight factor associated with the transportation cost;
- θ — weight factor associated with the inventory cost;
- z_α — standard normal deviate such that $P(z \le z_\alpha) = \alpha$;
- h — inventory holding cost per unit of product per year;
- $w_j(x)$ — total annual cost of working inventory held at distribution center j if the expected daily demand at j is x for each $j \in I$;
- F_j — fixed cost of placing an order at distribution center j, for each $j \in I$;
- L — lead time in days.

Decision Variables

$X_j = 1$, if retailer j is selected as a distribution center location, and 0 otherwise for each $j \in I$;

$Y_{ij} = 1$, if retailer i is served by a distribution center based at retailer j, and 0 otherwise for each $i \in I$ and each $j \in I$.

The model can now be formulated as follows:

$$\text{Minimize} \sum_{j\in I}\left\{ f_j X_j + \beta \sum_{i\in I} \mu_i d_{ij} Y_{ij} + w_j\left(\sum_{i\in I}\mu_i Y_{ij}\right) + \theta h z_\alpha \sqrt{\sum_{i\in I}\sigma_i^2 Y_{ij}}\right\}, \quad (1)$$

$$\text{subject to} \quad \sum_{j\in I} Y_{ij} = 1, \quad \text{for each } i \in I; \quad (2)$$

$$Y_{ij} - X_j \le 0, \quad \text{for each } i, j \in I; \quad (3)$$

$$Y_{ij} \in \{0, 1\}, \quad \text{for each } i, j \in I; \quad (4)$$

$$X_j \in \{0, 1\}, \quad \text{for each } j \in I. \quad (5)$$

The objective function (1) minimizes the weighted sum of the following four costs:

- the fixed cost of locating facilities,
- the shipping cost from the distribution centers to the non-DC retailers,
- the expected working inventory cost,
- the safety stock costs.

Constraint (2) stipulates that each retailer is assigned to exactly one distribution center. Constraint (3) states that retailers can only be assigned to candidate sites that are selected as distribution centers. Constraints (4) and (5) are standard integrality constraints.

The cost terms of the objective function warrant further explication. The first term is the single period (annual) amortization of the cost of establishing the chosen distribution centers. The second term represents the shipping cost from the distribution center(s) to the retailers. We assume that this cost is linear in the amount shipped. We further assume that shipments to the retailers occur on a scheduled basis with the frequency of shipments to each retailer determined exogenously. The third term represents the cost of the working inventory. This includes the fixed costs of placing orders as well as the shipment costs from the supplier to the distribution centers as well as the holding cost of the working inventory. For simplicity, we temporarily drop the index (j) on the distribution center and we let D be the total annual (expected) demand going through the distribution center and n be the number of shipments per year from the supplier. The total annual fixed, shipment, and inventory-carrying costs at the distribution center are then given by:

$$Fn+\beta v\left(\frac{D}{n}\right)n+\theta\frac{hD}{2n}. \tag{6}$$

The first term of (6) represents the total fixed cost of placing n orders per year. The second term represents the shipment cost $v(D/n)$ per shipment multiplied by the number of shipments per year. Note that D/n is the expected shipment size per shipment. Finally, the third term is the cost of the average working inventory. On average, there will be $D/2n$ items of inventory on hand at a cost of h per item per year. Taking the derivative of this expression with respect to n, the number of shipments per year, we obtain

$$\begin{aligned}F&+\beta v\left(\frac{D}{n}\right)-\beta nv'\left(\frac{D}{n}\right)\left(\frac{D}{n^2}\right)-\theta\frac{hD}{2n^2}\\&=F+\beta v\left(\frac{D}{n}\right)-\beta v'\left(\frac{D}{n}\right)\left(\frac{D}{n}\right)-\theta\frac{hD}{2n^2}=0.\end{aligned}\tag{7}$$

If $v(x)$ is linear in x (e.g., if $v(x)=g+ax$), then $v'(x)$ is a constant (e.g., a) and the expression above becomes

$$F+\beta g+\beta a\frac{D}{n}-\beta a\frac{D}{n}-\theta\frac{hD}{2n^2}=F+\beta g-\theta\frac{hD}{2n^2}=0. \tag{8}$$

Solving for n, we obtain $n=\sqrt{((\theta hD)/2(F+\beta g))}$. Substituting this into the total cost (6) we obtain

$$\begin{aligned}&F\sqrt{\frac{\theta hD}{2(F+\beta g)}}+\beta g\sqrt{\frac{\theta hD}{2(F+\beta g)}}+\beta aD+\theta\frac{hD}{2}\sqrt{\frac{2(F+\beta g)}{\theta hD}}\\&=\sqrt{\frac{\theta hD(F+\beta g)}{2}}+\beta aD+\sqrt{\frac{\theta hD(F+\beta g)}{2}}\\&=\sqrt{2\theta hD(F+\beta g)}+\beta aD.\end{aligned}\tag{9}$$

However recall that D is simply the annual demand through the distribution center or $\sum_{i\in I}\mu_i Y_{ij}$. Thus, in this case, which is the case we explore throughout the remainder of the paper, the objective function becomes:

$$\begin{aligned}&\text{Minimize}\\&\sum_{j\in I}\left\{f_jX_j+\beta\left(\sum_{i\in I}\mu_i d_{ij}Y_{ij}\right)+\sqrt{2\theta h(F_j+\beta g_j)}\sqrt{\sum_{i\in I}\mu_iY_{ij}}+\beta a_j\left(\sum_{i\in I}\mu_iY_{ij}\right)+\theta hz_\alpha\sqrt{\sum_{i\in I}\sigma_i^2Y_{ij}}\right\}\\&=\sum_{j\in I}\left\{f_jX_j+\left[\sum_{i\in I}(\beta\mu_id_{ij}Y_{ij}+\beta a_j\mu_i)Y_{ij}\right]+\sqrt{2\theta h(F_j+\beta g_j)}\sqrt{\sum_{i\in I}\mu_iY_{ij}}+\theta hz_\alpha\sqrt{\sum_{i\in I}\sigma_i^2Y_{ij}}\right\},\end{aligned}\tag{10}$$

where g_j and a_j have the obvious interpretation as being the fixed and per-unit shipping costs to distribution center j.

One further comment about the working-inventory-cost term is in order. Here we are actually estimating the expected working inventory cost using the cost associated with the expected quantity shipped. This estimate could be inappropriate, depending on the working-inventory-cost function w_j. It would be interesting to examine conditions under which this estimate is reasonable, and to examine further the cases in which this estimate is not reasonable. At this point we can say the following. If the function w_j is concave (as is the function we use in (7) assuming $v(x)$ is linear), then by Jensen's inequality (see Fergusen 1967, p. 77, for example) we know that $E[w_j(x)]\le w_j(E[x])$, where x is the quantity shipped. Thus, in the case in which w_j is concave, our objective function consistently overestimates the long-haul working-inventory cost. At this point, we believe this is a reasonable estimate for the case where w_j is concave and nondecreasing.

The fourth term in the objective function is the safety stock inventory-cost term. If each retailer maintains its own inventory (i.e., there is no risk pooling), then the inventory needed at retailer i to maintain a service level of α is $z_\alpha \sigma_i \sqrt{L_i}$, where L_i is the lead time in shipping to retailer i. If several retailers are served from a regional distribution center located at retailer j, then the combined variance at those retailers will be given by $\sum_{i\in I} \sigma_i^2 L_i Y_{ij}$ (assuming independence of the retailer demands). Thus, the inventory needed to maintain a service level of α is $z_\alpha \sqrt{\sum_{i\in I} \sigma_i^2 L_i Y_{ij}}$, assuming retailer demands are normally distributed. If all retailers experience the same lead time ($L_i = L$ for each $i \in I$), then the necessary inventory is $z_\alpha \sqrt{L} \sqrt{\sum_{i\in I} \sigma_i^2 Y_{ij}}$. To simplify the notation, we assume that all lead times are equal and that $\widehat{\sigma}_i^2$ is the lead-time variance at retailer i.

In summary, the objective function we consider is:

Minimize

$$\sum_{j\in I} \left\{ f_j X_j + \left[\sum_{i\in I} (\beta\mu_i d_{ij} Y_{ij} + \beta a_j \mu_i) Y_{ij} \right] + \sqrt{2\theta h (F_j + \beta g_j)} \sqrt{\sum_{i\in I} \mu_i Y_{ij}} + \theta h z_\alpha \sqrt{\sum_{i\in I} \widehat{\sigma}_i^2 Y_{ij}} \right\}$$

$$= \sum_{j\in I} \left\{ f_j X_j + \left(\sum_{i\in I} \widehat{d}_{ij} Y_{ij} \right) + K_j \sqrt{\sum_{i\in I} \mu_i Y_{ij}} + q \sqrt{\sum_{i\in I} \widehat{\sigma}_i^2 Y_{ij}} \right\}, \qquad (11)$$

where

$$\widehat{d}_{ij} = \beta\mu_i (d_{ij} + a_j);$$
$$K_j = \sqrt{2\theta h (F_j + \beta g_j)};$$
$$q = \theta h z_\alpha.$$

The constraints of the model are identical to those of the uncapacitated fixed-charge location problem. The first two terms of the objective function are structurally identical to those of the uncapacitated fixed-charge model. The remaining terms are nonlinear in the assignment variables. These terms significantly complicate the model and will, as a result, necessitate the development of solution algorithms that go beyond the algorithms for the standard uncapacitated fixed-charge problem. Known methods for the uncapacitated fixed-charge problem could, perhaps, be successfully modified, though we leave that approach for future research. Finally, we note that this model is, in some ways, similar to the "popular" mean/variance objective models of portfolio analysis; see Luenberger (1998).

4. The Set-Covering Model

4.1. Introduction

In this section, we formulate our decision problem as a set-covering model, and we present an approach to solving this model. The set-covering model has been intensely studied for at least 25 years. Ceria et al. (1997) give a nice historical survey. The widespread practical success of set-covering models for combinatorial optimization problems follows largely from the empirically typical tightness of the bound given by the associated linear-programming relaxations. This typical tightness appears in our situation as well.

A common approach to solving set-covering models is branch-and-price, a variant of branch-and-bound in which the nodes are processed by solving linear-programming relaxations via column-generation. For an introduction to branch-and-price, see Barnhart et al. (1994).

The set-covering model for our (nonlinear) location-inventory problem looks standard at first; it has no nonlinearity in its objective function or constraints. However, the nonlinearity appears in the objective function of the pricing problem, which must be solved to generate promising columns during the column-generation method. By exploiting particular special structures in two important special cases, we are able to solve our nonlinear pricing problem very efficiently. We thus have an effective method for processing the nodes during branch-and-price in these cases. We also show that the general pricing problem is polynomially solvable via the ellipsoid method.

We now present the set-covering model. First note that every solution to our decision problem consists of a partition of the set I of retailers into nonempty subsets, $R_1, R_2, \ldots, R_n$, together with one designated retailer for each of these n sets.

Let $\mathcal{R}$ be the collection of all nonempty subsets of the set I. For each set $R \in \mathcal{R}$, and each member $j \in R$, let $c_{R,j}$ be the total cost associated with establishing a distribution center at j and assigning this distribution center to serve the set R of retailers. That is,

$$c_{R,j} = f_j + \sum_{i \in R} \widehat{d}_{ij} + K_j \sqrt{\sum_{i \in R} \mu_i} + q \sqrt{\sum_{i \in R} \widehat{\sigma}_i^2}.$$

Now we define c_R to be the lowest cost of having one distribution center serve exactly the set R. That is, $c_R = \min_{j \in R} C_{R,j}$.

Note that here we assume a distribution center always serves itself. This may not be the case for the optimal solution to our decision problem, as stated. Indeed, in Appendix A we give an example in which some retailer is chosen as a distribution center location in the optimal solution but does not serve itself. The method we propose can be modified in a straightforward way to allow for this possibility.

The set-covering model has one variable for each set $R \in \mathcal{R}$:

$Z_R =$ 1 if set R is in the solution, and 0 otherwise for each $R \in \mathcal{R}$.

Now the model, which we will call $\mathcal{M}_{\mathcal{R}}$, can be expressed as follows:

$$\begin{aligned} \mathcal{M}_{\mathcal{R}}: \quad & \text{Minimize} \quad \sum_{R \in \mathcal{R}} c_R Z_R \\ & \text{subject to} \quad \sum_{R \in \mathcal{R}:\ i \in R} Z_R \ge 1, \quad \text{for each } i \in I; \\ & \qquad Z_R \in \{0, 1\}, \qquad \text{for each } R \in \mathcal{R}. \end{aligned}$$

Because the set $\mathcal{R}$ has exponentially many members, a branch-and-price approach is appropriate to solve this integer-programming problem. In branch-and-price applied to $\mathcal{M}_{\mathcal{R}}$, column generation is used to solve the linear-programming relaxations arising within a branch-and-bound approach. The next section details this step.

4.2. Solving the LP-Relaxation of the Set-Covering Model

Let us denote by $\overline{\mathcal{M}_{\mathcal{R}}}$ the linear-programming relaxation of $\mathcal{M}_{\mathcal{R}}$. To solve $\overline{\mathcal{M}_{\mathcal{R}}}$ via column generation, we start with a manageable subset $\mathcal{R}' \subset \mathcal{R}$. We will say more about generating an initial set $\mathcal{R}'$, but for now it is sufficient to take $\mathcal{R}'$ to be the collection of singleton subsets of I. Given $\mathcal{R}'$, the corresponding *master problem*, which we will call $\overline{\mathcal{M}_{\mathcal{R}'}}$, is:

$$\begin{aligned} \overline{\mathcal{M}_{\mathcal{R}'}}: \quad & \text{Minimize} \quad \sum_{R \in \mathcal{R}'} c_R Z_R \\ & \text{subject to} \quad \sum_{R \in \mathcal{R}':\ i \in R} Z_R \ge 1, \quad \text{for each } i \in I; \\ & \qquad 0 \le Z_R \le 1, \qquad \text{for each } R \in \mathcal{R}'. \end{aligned}$$

We begin each iteration by solving $\overline{\mathcal{M}_{\mathcal{R}'}}$, obtaining an optimal solution $\bar{Z}_R$, $R \in \mathcal{R}'$, and corresponding optimal dual solution $\bar{\pi}_i$, $i \in I$. Note that $\bar{Z}$ can be extended to a feasible solution to $\overline{\mathcal{M}_{\mathcal{R}}}$ by setting $\bar{Z}_R = 0$, for each $R \in \mathcal{R} \backslash \mathcal{R}'$. In this way, we will view $\bar{Z}$ as a feasible solution to $\overline{\mathcal{M}_{\mathcal{R}}}$.

Now $\bar{Z}$ is an optimal solution to $\overline{\mathcal{M}_{\mathcal{R}}}$, if every set $R \in \mathcal{R}$ has nonnegative reduced cost, with respect to $\bar{\pi}$. That is,

$$c_R - \sum_{i \in R} \bar{\pi}_i \ge 0,$$

for each $R \in \mathcal{R}$. If, on the other hand, a set R with negative reduced cost is found, then R is added to $\mathcal{R}'$, and the next iteration begins.

Finding $R \subset I$ with negative reduced cost, or proving that no such R exists, is called the *pricing problem*. To solve the pricing problem for our model, it suffices to find, for every retailer $j \in I$, a minimum-reduced-cost set $R_j^* \subset I$, say, having j as the designated distribution center. If every R_j^* has nonnegative reduced cost, then every $R \subset I$ has nonnegative reduced cost.

Thus, the pricing problem reduces to finding R_j^*, for each $j \in I$. To find R_j^* we must solve the following integer-programming problem, $\mathcal{P}_j$:

$$\begin{aligned} \mathcal{P}_j: \quad & \text{Minimize} \quad f_j + \sum_{i \in I} (\widehat{d}_{ij} - \bar{\pi}_i) Y_{ij} + K_j \sqrt{\sum_{i \in I} \mu_i Y_{ij}} \\ & \qquad + q \sqrt{\sum_{i \in I} \widehat{\sigma}_i^2 Y_{ij}} \\ & \text{subject to} \quad Y_{ij} \in \{0, 1\}, \quad \text{for each } i \in I; \\ & \qquad Y_{jj} = 1. \end{aligned}$$

Given an optimal solution Y^* to $\mathcal{P}_j$, the set R_j^* is then the set $\{i \in I: Y_{ij}^* = 1\}$.

In what follows, we discuss solutions to two special cases of the pricing problem $\mathcal{P}_j$. The general case is handled in Appendix B.

First Special Case: The Pricing Problem When the Variance of Demand Is Proportional to the Mean. We now show how to solve the pricing problem when the variance of demand of retailer i is proportional to the mean demand of retailer i and that the proportionality constant is the same for each retailer i. In other words, we assume $\mu_i = \gamma \sigma_i^2 \ \forall i \in I$. Thus, we are finding R_j^* for designated distribution center j under this assumption. Under this assumption (which is reasonable for many demand distributions, including the Possion distribution), the objective function of $\mathcal{P}_j$ becomes

$$f_j + \sum_{i\in I} (\widehat{d}_{ij} - \bar{\pi}_i) Y_{ij} + K_j \sqrt{\sum_{i\in I} \gamma \sigma_i^2 Y_{ij}} + q\sqrt{\sum_{i\in I} \widehat{\sigma}_i^2 Y_{ij}}$$
$$= f_j + \sum_{i\in I} (\widehat{d}_{ij} - \bar{\pi}_i) Y_{ij} + \left(K_j\sqrt{\gamma} + q\sqrt{L}\right) + \sqrt{\sum_{i\in I} \widehat{\sigma}_i^2 Y_{ij}}.$$

To simplify the notation, we define

$$b_i = \widehat{d}_{ij} - \overline{\pi}_i,$$
$$g_i := \sigma_i^2 \left(K_j\sqrt{\gamma} + q\sqrt{L}\right)^2, \quad \text{and}$$
$$y_i = Y_{ij},$$

for each $i \in I$. We now have the following pricing problem $\mathcal{P}_j$, for designated distribution center $j \in I$:

$$\mathcal{P}_j: \quad \text{Minimize} \quad \sum_{i\in I} b_i y_i + \sqrt{\sum_{i\in I} g_i y_i}$$
$$\text{subject to} \quad y_i \in \{0, 1\}, \quad \text{for each } i \in I;$$
$$y_j = 1.$$

Let y^* be an optimal solution to $\mathcal{P}_j$, with associated optimal objective value w_j^*. The minimum-reduced-cost set $R_j^* \subset I$ is then the collection of retailers $i \in I$ with $y_i^* = 1$. If $w_j^* + f_j \geq 0$, then R_j^* has nonegative reduced cost; moreover, we can conclude that there is no set $R \in \mathcal{R}$ having designated distribution center j with negative reduced cost. Further, if, for each $j \in I$, we find that $w_j^* + f_j \geq 0$, then we can conclude that there is no set $R \in \mathcal{R}$ with negative reduced cost.

We begin the analysis of problem $\mathcal{P}_j$ with an observation.

Lemma 4.1. *Given a retailer $j \in I$, and associated minimum-reduced-cost set $R_j^* \subset I$, for every $i \in R_j^* \backslash \{j\}$, $b_i \leq 0$.*

Proof. Let $i \in R_j^* \backslash \{j\}$. Since $g_i \geq 0$, if $b_i > 0$, then for any solution $\bar{y}$ with $\bar{y}_i = 1$ the objective function value is strictly greater than that of the solution obtained from $\bar{y}$ by setting $\bar{y}_i = 0$. □

To find R_j^*, we may therefore restrict our attention to the retailers $I^- := \{i \in I \backslash \{j\}: b_i \leq 0\}$. Now assume that this subset I^- of retailers has been sorted so that:

$$g_1 > 0, \quad g_2 > 0, \dots, \quad g_{m_1} > 0,$$
$$g_{m_1+1} = g_{m_1+2} = \cdots = g_{m_1+m_2} = 0,$$

and

$$\frac{b_1}{g_1} \leq \frac{b_2}{g_2} \leq \cdots \leq \frac{b_{m_1}}{g_{m_1}} \leq 0.$$

Theorem 4.2. *There is an optimal solution y^* to problem $\mathcal{P}_j$, in which the following four properties hold*:

(1) *$y_i^* = 0$, for each $i \in I \backslash (I^- \cup \{j\})$—that is, $y_i^* = 0$ if $b_i > 0$);*

(2) *$y_j^* = 1$;*

(3) *$y_{m_1+1}^* = y_{m_1+2}^* = \cdots = y_{m_1+m_2}^* = 1$—that is, all retailers i with $b_i \leq 0$ that have zero demand variance are served from j;*

(4) *if $y_k^* = 1$ for some $k \in \{1, \dots, m_1\}$, then $y_l^* = 1$ for all $l \in \{1, 2, \dots, k-1\}$.*

Proof. Property 1 follows from Lemma 4.1. Property 2 follows immediately from feasibility of y^*. Property 3 follows because for any retailer i with $g_i = 0$ and $b_i \leq 0$, setting $y_i^* = 1$ maintains feasibility and does not increase the objective value.

Now assume we have an optimal solution y^* satisfying Properties 1, 2, and 3, and assume $y_k^* = 1$, for some $k \in \{1, \dots, m_1\}$. Suppose, for a contradiction, that $y_l^* = 0$, for some $l \in \{1, 2, \dots, k-1\}$.

Define two new solutions y' and y'' as follows:

$$y_i' = \begin{cases} 1, & \text{if } i = l \\ y_i^*, & \text{otherwise,} \end{cases}$$

and

$$y_i'' = \begin{cases} 0, & \text{if } i = k \\ y_i^*, & \text{otherwise,} \end{cases}$$

for each $i \in I$.

Let z^*, z', z'' be the objective values of y^*, y', y'', respectively. We will show that because $z^* - z'' \le 0$, by optimality of y^*, it follows that $z' - z^* \le 0$. From this it follows that y' is also an optimal solution. Thus, if we take at the outset y^* to be an optimal solution satisfying Properties 1, 2, and 3 with the largest number of the variables $y_1, y_2, \ldots, y_{m_1}$ set to 1, then it follows from this argument that y^* satisfies Condition 4, as desired.

It is immediate that

$$z' - z^* = b_l + \sqrt{\sum_{i\in R} g_i + g_k + g_l} - \sqrt{\sum_{i\in R} g_i + g_k}$$

and

$$z^* - z'' = b_k + \sqrt{\sum_{i\in R} g_i + g_k} - \sqrt{\sum_{i\in R} g_i}$$

for an appropriately defined $R \subset I$. For notation simplicity, let $\overline{G} = \sum_{i\in R} g_i$. Then

$$z' - z^* = b_l + \sqrt{\overline{G} + g_k + g_l} - \sqrt{\overline{G} + g_k}$$

and

$$z^* - z'' = b_k + \sqrt{\overline{G} + g_k} - \sqrt{\overline{G}}.$$

Since $(b_l/g_l) \le (b_k/g_k)$, by assumption, and

$$\frac{\sqrt{\overline{G} + g_k + g_l} - \sqrt{\overline{G} + g_k}}{g_l} \le \frac{\sqrt{\overline{G} + g_k} - \sqrt{\overline{G}}}{g_k},$$

by concavity of the square-root function, we have that

$$\frac{z' - z^*}{g_l} = \frac{b_l}{g_l} + \frac{\sqrt{\overline{G} + g_k + g_l} - \sqrt{\overline{G} + g_k}}{g_l}$$

$$\le \frac{b_k}{g_k} + \frac{\sqrt{\overline{G} + g_k} - \sqrt{\overline{G}}}{g_k} = \frac{z^* - z''}{g_k}.$$

Because $g_k > 0$ and $g_l > 0$, it follows that since $z^* - z'' \le 0$ (because y^* is optimal), we have that $z' - z^* \le 0$, as desired. □

Theorem 4.2 provides an efficient method for solving the pricing problem $\mathcal{P}_j$, namely, we simply generate all solutions with Properties 1, 2, 3, and 4, and then select the one with the lowest objective value. It is easy to see there are at most $|I|$ such solutions. They can be listed immediately after sorting the terms b_i/g_i. With a bit of reflection, we see we can compute the entire list of objective functions with only $O(|I|)$ multiplications, additions, and square-root computations. Thus, sorting, which can be done in time $O(|I|\log(|I|))$, is the dominant effort. Therefore, solving the pricing problem, which involves solving $\mathcal{P}_j$ for each $j \in I$ can be done in time $O(|I|^2 \log(|I|))$.

Second Special Case: The Pricing Problem When Demand Has Zero Variance. In this part, we show how to solve the pricing problem that arises from the set-covering model associated with the original objective function (1) under the condition that the demand variance is zero, for every retailer. That is, we find R_j^* for designated distribution center j, assuming $\sigma_i = 0$, for every $i \in I$. We do this for the case of a general working-inventory-cost function $w_j(x)$ as long as the function is nonnegative and concave in the demand assigned to distribution center j. In particular, in this subsection we temporarily drop the assumption that the shipment cost to distribution center j ($v_j(x)$) is linear in the demand assigned to distribution center j. Thus, we are clustering retailers to save on the working inventory (and possibly the plant to distribution center shipment) costs rather than to achieve any risk-pooling benefits.

In this case, the objective function of the pricing problem associated with distribution center j becomes

$$\text{Minimize } f_j + \sum_{i\in I} (\beta\mu_i d_{ij} - \bar{\pi}_i) Y_{ij} + \theta w_j\left(\sum_{i\in I} \mu_i Y_{ij}\right).$$

To simplify the notation, we again eliminate the constant term f_j and make the following substitutions

$$b_i = \beta\mu_i d_{ij} - \overline{\pi_i},$$

$$y_i = Y_{ij},$$

for each $i \in I$, and

$$\eta(x) = \theta w_j(x).$$

We now have the following pricing problem $\mathcal{P}_j$, for designated distribution center $j \in I$:

$$\mathcal{P}_j: \quad \text{Minimize} \quad \sum_{i\in I} b_i y_i + \eta\left(\sum_{i\in I} \mu_i y_i\right)$$

$$\text{subject to} \quad y_i \in \{0, 1\}, \quad \text{for each } i \in I;$$

$$y_j = 1.$$

Let y^* be an optimal solution to $\mathscr{P}_j$, with associated optimal objective value w_j^*. The minimum-reduced-cost set $R_j^* \subset I$ is then the collection of retailers $i \in I$ with $y_i^* = 1$.

We begin the analysis of problem $\mathscr{P}_j$ with an observation.

Lemma 4.3. *Given a retailer $j \in I$, and associated minimum-reduced-cost set $R_j^* \subset I$, for every $i \in R_j^* \backslash \{j\}$, $b_i \leq 0$.*

Proof. Since $\eta(x)$ is monotone nondecreasing, if $b_i > 0$, for some $i \in R_j^* \backslash \{j\}$, then for any solution $\bar{y}$ with $\bar{y}_i = 1$, the objective function value is strictly greater than that of the solution obtained from $\bar{y}$ by setting $\bar{y}_i = 0$. □

To find R_j^*, we may therefore restrict our attention to the retailers $I^- := \{i \in I \backslash \{j\} \ : \ b_i \leq 0\}$. Now assume that this subset I^- of retailers has been sorted so that:

$$\frac{b_1}{\mu_1} \leq \frac{b_2}{\mu_2} \leq \cdots \leq \frac{b_m}{\mu_m} \leq 0.$$

Theorem 4.4. *There is an optimal solution y^* to $\mathscr{P}_j$ in which the following three properties hold*:

(1) *$y_i^* = 0$, for each $i \in I \backslash (I^- \cup \{j\})$—that is, $y_i^* = 0$ if $b_i > 0$;*

(2) *$y_j^* = 1$;*

(3) *if $y_k^* = 1$, for some $k \in \{1, \ldots, m\}$, then $y_l^* = 1$, for all $l \in \{1, 2, \ldots, k-1\}$.*

Proof. Property 1 follows from Lemma 4.3. Property 2 follows immediately from feasibility of y^*.

Now assume we have an optimal solution y^* satisfying Properties 1 and 2, and assume $y_k^* = 1$ for some $k \in \{1, \ldots, m\}$. Suppose, for a contradiction, that $y_l^* = 0$, for some $l \in \{1, 2, \ldots, k-1\}$.

Define two new solutions y' and y'' as follows:

$$y'_i = \begin{cases} 1, & \text{if } i = l \\ y_i^*, & \text{otherwise,} \end{cases}$$

and

$$y''_i = \begin{cases} 0, & \text{if } i = k \\ y_i^*, & \text{otherwise,} \end{cases}$$

for each $i \in I$.

Let z^*, z', z'' be the objective values of y^*, y', y'', respectively. We will show that because $z^* - z'' \leq 0$, by optimality of y^*, it follows that $z' - z^* \leq 0$. From this it follows that y' is also an optimal solution. Thus, if we take at the outset y^* to be an optimal solution satisfying Properties 1 and 2 with the largest number of the variables $y_1, y_2, \ldots, y_m$ set to 1, then it follows from this argument that y^* satisfies Property 3, as desired.

It is immediate that

$$z' - z^* = b_l + \eta\left(\sum_{i\in R}\mu_i + \mu_k + \mu_l\right) - \eta\left(\sum_{i\in R}\mu_i + \mu_k\right)$$

and

$$z^* - z'' = b_k + \eta\left(\sum_{i\in R}\mu_i + \mu_k\right) - \eta\left(\sum_{i\in R}\mu_i\right)$$

for an appropriately defined $R \subset I$. For notation simplicity, define $a = \sum_{i\in R}\mu_i$. Then

$$z' - z^* = b_l + \eta\,(a + \mu_k + \mu_l) - \eta\,(a + \mu_k)$$

and

$$z^* - z'' = b_k + \eta\,(a + \mu_k) - \eta\,(a)\,.$$

Now, since $(b_l/\mu_l) \leq (b_k/\mu_k)$, and, by concavity of the function η,

$$\frac{\eta\,(a+\mu_k+\mu_l) - \eta\,(a+\mu_k)}{\mu_l} \leq \frac{\eta\,(a+\mu_k) - \eta\,(a)}{\mu_k},$$

we have that

$$\begin{aligned} \frac{z' - z^*}{\mu_l} &= \frac{b_l}{\mu_l} + \frac{\eta\,(a+\mu_k+\mu_l) - \eta\,(a+\mu_k)}{\mu_l} \\ &\leq \frac{b_k}{\mu_k} + \frac{\eta\,(a+\mu_k) - \eta\,(a)}{\mu_k} = \frac{z^* - z''}{\mu_k}. \end{aligned}$$

Since $\mu_k > 0$ and $\mu_l > 0$, it follows that since $z^* - z'' \leq 0$, we have that $z' - z^* \leq 0$, as desired. □

As in the first special case, the pricing problem in this case can be solved in time $O(|I|^2 \log(|I|))$.

5. Computational Results

In this section, we summarize our computational experience with the column-generation method outlined in the previous section.

We generated a total of 47 instances using four different data sets: (1) a 33-city problem found in Nemhauser and Wolsey (1988); (2) a 49-city problem; (3) an 88-city problem; and (4) a 150-city problem (2, 3, and 4 are all from Daskin 1995). For Data Set (1), we randomly generated the mean demands μ_i, ranging from 100 to 1,600. For each of the four data sets, we set variances $\sigma_i^2 = \mu_i$, for all $i \in I$. We set holding cost to be 1, $z_\alpha = 1.96$ (97.5% service level), $a_i = 5$, $g_i = 10$, and $F_i = 10$ for all $i \in I$. Then we generated instances by varying the values of β (the distribution cost factor) and θ (the inventory holding cost factor). Our goal was to find ranges of values for β and θ that resulted in instances that varied in solution difficulty as well as the fraction of retailers used as distribution centers in the solution.

Tables 1, 2, 3, and 4 highlight the results of our computational study. For each of the instances, we first solved the linear-programming relaxation of the set-covering model $\mathcal{M}_{\mathcal{R}}$ via column generation. The initial set of columns included all singletons, together with the columns corresponding to the optimal solution to the Uncapacitated Facility Location Problem for the instance. When solving the pricing problem $\mathcal{P}_j$, we solved it for every j and added all the columns with negative reduced cost.

The column labeled *Number of Iterations* indicates the number of pricing problems that were solved during this phase. The column labeled *Number of Columns Generated* indicates the total number of columns added during this phase. The column labeled "Number of DCs Opened" indicates the number of sets R with value $Z_R^* = 1$ in the optimal linear-programming solution. In all instances we tested, the corresponding optimal solution we obtained was integral.

The algorithm is coded in C++, and the linear- and integer-programming problems were solved using CPLEX 5.0. The instances were solved on a Sun SPARCstation running the SunOS 4.1.3u5 operating system. All computation times exclude input times.

Table 1 Computation Results for 33-City Problem

INPUT		OUTPUT				
β	θ	Number of Iterations	Number of DCs Opened	CPU Time (Seconds)	Number of Columns Generated	Comments
0.001	0.1	33	4	33	2,206	
0.002	0.1	15	6	6	1,002	
0.003	0.1	8	7	1	427	
0.004	0.1	9	8	1	372	
0.005	0.1	8	10	1	283	
0.001	0.1	33	4	33	2,206	
0.002	0.2	14	6	5	1,029	
0.005	0.5	8	9	1	287	
0.005	0.1	8	10	1	283	
0.005	1	9	9	1	307	
0.005	5	6	8	1	348	
0.005	10	9	7	1	427	
0.005	20	11	6	3	671	

Table 2 Computation Results for 49-City Problem

INPUT		OUTPUT				
β	θ	Number of Iterations	Number of DCs Opened	CPU Time (Seconds)	Number of Columns Generated	Comments
0.001	1	66	5	102	4,217	1, 3, 5, 6, 22
0.002	1	24	7	12	1,656	1, 2, 3, 5, 7
0.003	1	13	10	3	662	22, 30
0.001	1	66	5	102	4,217	
0.005	5	8	15	1	459	
0.01	10	10	16	3	928	
0.005	5	8	15	1	459	
0.005	10	14	19	6	1,314	
0.005	20	31	7	48	3,841	

Table 3 Computation Results for 88-City Problem

INPUT		OUTPUT				
β	θ	Number of Iterations	Number of DCs Opened	CPU Time Seconds	Number of Columns Generated	Comments
0.001	0.1	111	9	626	12,246	
0.002	0.1	50	11	65	12,246	
0.003	0.1	24	15	22	3,417	
0.004	0.1	24	21	15	1,409	
0.005	0.1	18	23	8	846	
0.001	0.1	111	9	626	12,246	
0.002	0.2	56	10	89	4,185	
0.005	0.5	14	22	7	9,016	
0.005	0.1	18	23	8	846	
0.005	0.5	14	22	7	906	
0.005	1	21	21	12	1,222	
0.005	5	32	17	60	4,579	
0.005	10	34	12	145	9,158	
0.005	20	66	9	869	20,114	

Table 4 Computation Results for 150-City Problem

INPUT		OUTPUT				
β	θ	Number of Iterations	Number of DCs Opened	CPU Time Seconds	Number of Columns Generated	Comments
0.0004	0.01	55	15	252	9624	
0.0006	0.01	47	21	180	8075	
0.0008	0.01	30	36	55	3351	
0.001	0.01	20	28	29	2318	
0.0005	0.01	54	18	277	10907	
0.001	0.02	28	28	45	2730	
0.002	0.04	19	41	22	1286	
0.001	0.01	20	28	29	2318	
0.001	0.1	45	26	102	4233	
0.001	0.5	44	21	261	11795	
0.001	1	57	15	730	21944	

The initial columns from solving the corresponding uncapacitated facility location problem are given as input for all instances.

Table 2 presents the results of the 49-city problem. When $\beta = 0.001$ and $\theta = 1$, there are five cities that serve as distribution centers in the optimal solution. They are Cities 1, 3, 5, 6, and 22. After we increase β to 0.002, the number of distribution centers increases to seven. They are Cities 1, 2, 3, 5, 7, 22, and 30. As we can see here, those five cities are not a subset of these seven cities. Thus, greedy algorithms that myopically add distribution centers to the solution are not likely to be very effective.

As we can see from these four tables, the problem is difficult to solve when β (the distribution cost factor) is small or θ (the inventory holding cost factor) is large. When we increase the value of θ while fixing the value of β, the solution time increases, the number of columns generated increases, and the number of iterations needed to solve the problem tends to increase. Similarly, when we fix the value of θ and decrease the value of β, we can observe the same trends. For example, for the 88-city problem, Figure 1 illustrates the relationship between CPU time and the value of θ when we fix $\beta = 0.005$, and Figure 2 illustrates the relationship between CPU time and the value of β when we fix $\theta = 0.1$. This is because θ is associated with the nonlinear term while β is mainly

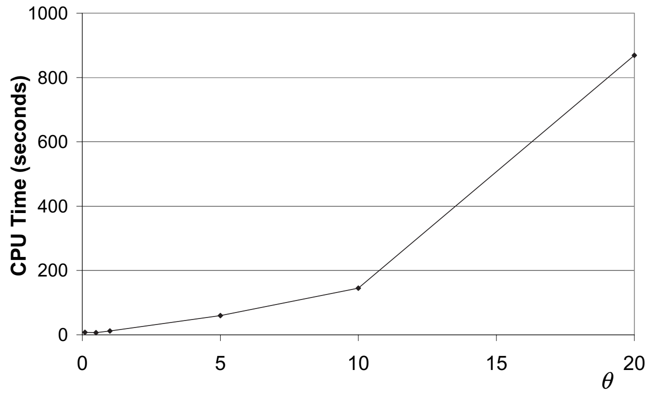

Figure 1 Computational Time Versus θ (88-City, $\beta = 0.005$)

associated with the linear term in the objective function. So, when the inventory decisions are crucial (θ/β is large), the problem becomes more nonlinear, so it is relatively more difficult to solve.

With regard to the number of distribution centers opened, we make the following observations. First, when β increases (i.e., the transportation costs increase relative to other costs), the number of opened distribution centers also increases. As we observed in Table 2, the smaller set of opened DCs is not necessarily a subset of the larger set of opened DCs. Second, when θ increases (i.e., the inventory costs increase relative to other costs) the number of opened distribution centers decreases. Again, for this case, we notice that the smaller set of opened DCs is not necessarily a subset of the larger set of opened DCs. Figure 3 shows the number of distribution centers opened with

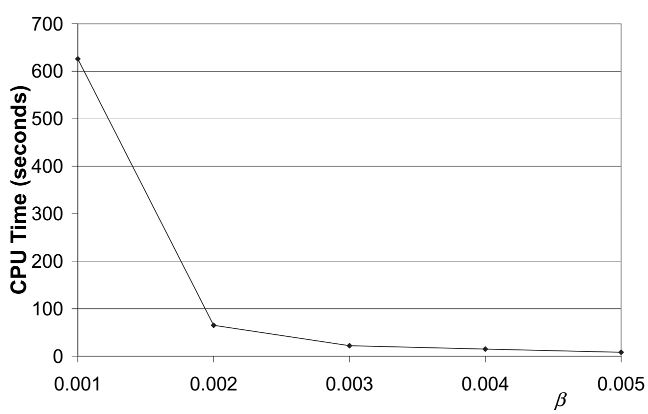

Figure 2 Computational Time Versus β (88-City, $\theta = 0.1$)

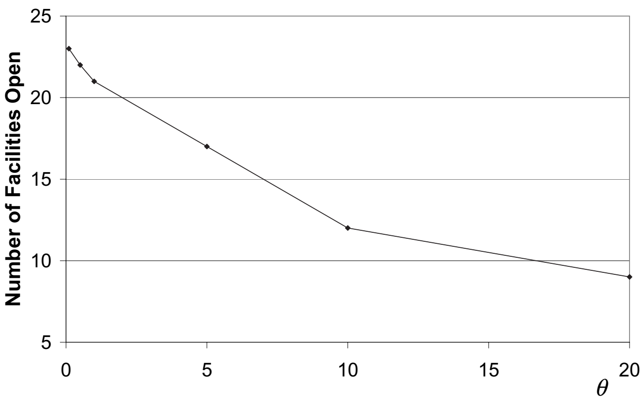

Figure 3 **Number of Distribution Centers Open Versus θ (88-City, $\beta = 0.005$)**

respect to the value of θ when we fix $\beta = 0.005$ in the 88-city problem. Figure 4 shows the number of distribution centers opened with respect to the value of β when $\theta = 0.1$ for the same problem.

We can also see from all these four tables, when we increase both β and θ (the inventory costs and the transportation costs increase relative to fixed costs), the number of opened distribution centers increases.

6. Conclusions and Future Research

In this paper we have outlined a formulation of an integrated facility location/inventory location model. The model determines the location of distribution centers and the assignment of retailers to the distribution centers to minimize the total fixed distribution-center location costs, working-inventory costs at the distribution centers which includes the line-haul shipment costs from a single supplier to the distribution centers, transport costs from the distribution center to the retailers, and safety stock inventory costs at the distribution centers. The formulation allows for the incorporation of nonlinear working-inventory costs and nonlinear safety stock inventory costs.

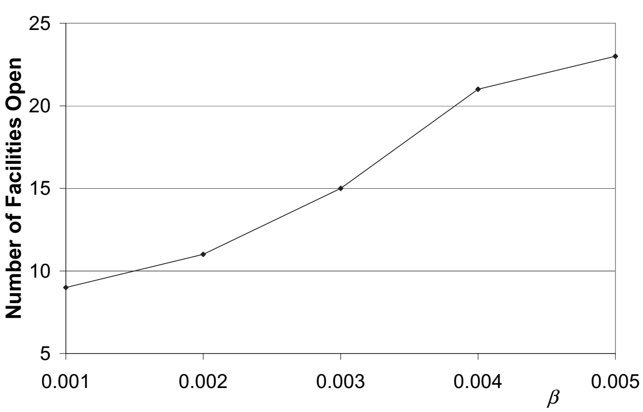

Figure 4 **Number of Distribution Centers Open Versus β (88-City, $\theta = 0.1$)**

The model was initially formulated as a mixed-integer nonlinear location allocation model. As such, it can be viewed as a direct extension of the classic uncapacitated fixed-charge location model. We then reformulated the problem as a linear integer set-covering problem. Two issues arise in this transformation. First, the number of columns required in the set-covering model is exponentially large. We attack this problem using column-generation techniques. This leads to the second issue. The pricing problem that must be solved at each iteration of the column-generation algorithm is nonlinear. However, we show that for two important cases—the case of linear plant to distribution-center transport costs coupled with the mean demand proportional to the variance of demand of each retailer, and the case of known retailer demands—the nonlinear problem can be solved exactly in polynomial time. The dominant step in the solution algorithm is sorting a list whose length equals the number of retailer nodes.

Computational results were provided for problems ranging in size from 33 nodes to 150 nodes and the variance of demand is proportional to the mean. The results suggest that as the nonlinear safety stock costs increase relative to the other costs (or equivalently as the desired service level improves), the problem becomes harder to solve. Also, as these costs increase, the number of distribution centers located decreases so that the system can make better use of the risk-pooling effect.

This research can be extended in a number of important ways. First, it would be important to identify other special cases of the problem that can be readily solved using the set-covering approach. It would be particularly important to see if methods could be devised to solve the problem with concave transport costs from the supplier to the distribution centers, or when the mean and variance of the retailer demands are not proportional to each other.

Second, inventory at the retailers has not been explicitly modeled. We have assumed that the retailers carry little, if any, inventory and that the shipment frequency to the retailers (and hence, their average inventory) is determined exogenously. This was appropriate in the context of the blood bank problem that motivated this model, but may not be appropriate in more general contexts. Thus, explicit incorporation and modeling of the shipment process from the distribution centers to the retailers, along with the retailer inventory management policies, would be desirable.

Third, we have not explicitly modeled the process of emergency shipments from the distribution center to the retailer in the event of a stockout at a retailer. This process should also be incorporated into the model. It is likely that this will lead to a stochastic programming formulation in which recourse actions are explicitly modeled.

Finally, because our computational results did not find any instances in which the problem could not be solved to optimality at the root node of the branch-and-bound tree, we have not implemented a full branch-and-price code. This too would be useful, particularly if additional computational tests suggest that the likelihood of finding noninteger solutions at the root node is high for some data sets or values of the parameters.

The paper has outlined a new model that integrates facility location and inventory risk-pooling effects explicitly. As such, we believe it is a valuable extension to the current inventory literature that treats facility locations as given and the facility location literature that tends to ignore inventory effects. It is, however, only a first step. Important additional research can and should flow from this work.

Acknowledgments
This work was supported in part by National Science Foundation Grants DMI-9523048 and DMI-9634750. The authors thank the anonymous referees for their valuable comments.

Appendix A. Example Where Distribution Center Does Not Serve Itself

In our computational study, we assumed that any retailer used as a distribution center will serve itself. We believe that in practice

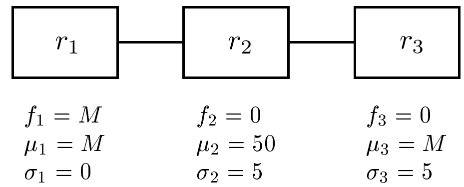

Figure 5 Example Where a Distribution Center Does Not Serve Itself

this would always be the case. In this section we present a class of examples to illustrate that this assumption can yield suboptimal solutions.

Consider a system in which the set of retailers is $I = \{r_1, r_2, r_3\}$, positioned on a line as shown in Figure 1. Assume the shipping cost from the supplier is the same for all feasible solutions and can therefore be ignored.

The situation in this example is as follows. Retailer r_1 has a prohibitively high fixed cost f_1 and will therefore not be a distribution center in any optimal solution. Retailer r_1 also has significantly high mean demand μ_1, with no variance, $\sigma_1 = 0$. Retailers r_2 and r_3 have a fixed cost of zero, and so they may both be distribution centers in the optimal solution. It is cheaper to ship from r_2 to r_1 than it is to ship from r_3 to r_1. Therefore, because r_1 has no demand variance, r_2 supplies r_1 in the optimal solution. Retailer r_3 has high demand μ_3 and high variance σ_3. Retailer r_2 has low demand μ_2 and high variance σ_2. There is benefit from pooling these two demands, because $\sqrt{\sigma_2^2} + \sqrt{\sigma_3^2} > \sqrt{\sigma_2^2 + \sigma_3^2}$. But μ_3 is prohibitively high, and therefore transportation costs are least if r_3 supplies r_2. Thus, r_2 will supply r_1, but r_3 will supply r_2, provided the risk-pooling benefit outweighs the shipping cost.

To make this formal and concrete, let us take the unit shipping costs from r_2 to r_3 and from r_3 to r_2 to be 1 and let G represent the constant $\theta h z_\alpha$ in the inventory holding cost term. Now r_3 will supply r_2 in the optimal solution, provided that

$$G\sigma_2 + G\sigma_3 > \mu_2 + G\sqrt{\sigma_2^2 + \sigma_3^2}.$$

This is the case whenever

$$G > \frac{\mu_2}{\sigma_2 + \sigma_3 - \sqrt{\sigma_2^2 + \sigma_3^2}}.$$

For a numerical example, let $\sigma_2 = \sigma_3 = 5$ and let $\mu_2 = 50$. Then r_3 will supply r_2 whenever $G > 17.07$.

Appendix B. Solving the General Pricing Problem in Polynomial Time

In this appendix, we show that assuming only that the working-inventory cost function $w_j(x)$ is concave, the general pricing problem can be restated as the problem of minimizing a particular submodular function, and can therefore be solved in polynomial time. We do this by showing that the objective function of the pricing problem $\mathcal{P}_j$ is submodular on the set $E = I \setminus \{j\}$.

Recall our set-covering model $\mathcal{M}_{\mathcal{R}}$.

$$\mathcal{M}_{\mathcal{R}}: \quad \text{Minimize} \quad \sum_{R\in\mathcal{R}} c_R Z_R$$
$$\text{subject to} \quad \sum_{R\in\mathcal{R}: i\in R} Z_R \geq 1, \quad \text{for each } i \in I;$$
$$Z_R \in \{0,1\}, \quad \text{for each } R \in \mathcal{R}.$$

In its most general form, without the linearity assumption on the shipping cost function $v_j(x)$, the cost coefficient that results from (1) is $c_R = \min_{j\in R} C_{R,j}$, where

$$c_{R,j} = f_j + \beta \sum_{i\in R} \mu_i d_{ij} + w_j\left(\sum_{i\in R}\mu_i\right) + \theta h z_\alpha \sqrt{\sum_{i\in R}\sigma_i^2}.$$

In this appendix, we assume only that the working-inventory cost function $w_j(x)$ is concave. Now the pricing problem $\mathcal{P}_j$ is:

$$\mathcal{P}_j: \quad \text{Minimize} \quad f_j + \sum_{i\in I}(\beta\mu_i d_{ij} - \bar{\pi}_i)Y_{ij} + w_j\left(\sum_{i\in I}\mu_i Y_{ij}\right) + \theta z_\alpha \sqrt{\sum_{i\in I}\sigma_i^2 Y_{ij}}$$
$$\text{subject to} \quad Y_{i,j} \in \{0,1\}, \quad \text{for each } i \in I;$$
$$Y_{j,j} = 1$$

The main theorem of this appendix is that this pricing problem is a submodular function minimization problem. We begin with a brief review of definitions and background on submodular functions.

Given a finite set E, a real-valued function $g(\cdot)$ that is defined on the subsets of E is called *submodular* if, for every pair $S, T \subseteq E$, we have that

$$g(S) + g(T) \geq g(S\cap T) + g(S\cup T).$$

Submodular functions have been studied as early as Rado (1942). Later, Edmonds (1970) renewed attention to submodular functions by proving the integrality of *polymatroids*, which are polytopes that arise from rational submodular functions. Edmonds' results generalized many well-known theorems about integral polytopes.

Submodular function minimization is the following problem. Given a submodular function $g(\cdot)$ defined on the subsets of a finite set E, find a subset $A^* \subseteq E$ such that $g(A^*) = \min\{g(A): A \subseteq E\}$. Minimizing a rational submodular function was shown by Grötschel et al. (1981) to be solvable in polynomial time via the ellipsoid algorithm. More recently, Iwata et al. (1999) and, independently, Schrijver (1999), have developed strongly polynomial combinatorial algorithms for the submodular function minimization problem. Thus, the result we present implies the general pricing problem is polynomially solvable. Moreover, because the submodular function yielded by our pricing problem is quite special, there is hope for an even more practical algorithm.

We begin with some elementary properties of submodular functions, the proofs of which are straightforward (See, for example, Lovasz 1982 or Fugishige 1991). For convenience, given a vector $d \in \Re^E$ and set $S \subset E$, define $d(S) = \sum_{i\in S} d_i$. In this way, d determines a function defined on the subsets of E, called *the set function defined by d*.

Property B.1. *If $d \in \Re^E$, then the set function defined by d is submodular.*

Property B.2. *Let g be a submodular function on set E, and let c be a nonnegative real number. For each $S \subseteq E$, define $cg(S) = c\cdot g(S)$. Then cg is a submodular function.*

Property B.3. *Let g be a submodular function on set E, and let c be a real number. For each $S \subseteq E$, define $(c+g)(S) = c + g(S)$. Then $c+g$ is a submodular function.*

Property B.4. *Let g and h be submodular functions on the same set E. For each $S \subseteq E$, define $(g+h)(S) = g(S) + h(S)$. Then $g+h$ is a submodular function.*

The next lemma shows how any concave function, together with a nonnegative vector, gives rise to a submodular set function.

Lemma B.1. *Let $v: \Re \to \Re$ be a concave function, and let $d \in \Re^E$ be a nonnegative vector. Define $g(S) = v(d(S))$, for each $S \subseteq E$. Then g is submodular on E.*

Proof. Let $S, T \subseteq E$ be arbitrary. If $S \subseteq T$, then the result is immediate. Thus, we can assume $S\setminus T \neq \emptyset \neq T\setminus S$. Next, note that since $d \geq 0$, we know that $d(S\cap T) \leq d(S) \leq d(S\cup T)$ and $d(S\cap T) \leq d(T) \leq d(S\cup T)$. Now, by the concavity of v:

$$\begin{aligned} v(d(S)) &= v(d(S\cap T)) + \frac{v(d(S)) - v(d(S\cap T))}{d(S) - d(S\cap T)}[d(S) - d(S\cap T)] \\ &\geq v(d(S\cap T)) + \frac{v(d(S\cup T)) - v(d(S\cap T))}{d(S\cup T) - d(S\cap T)}[d(S) - d(S\cap T)], \end{aligned}$$

and, similarly:

$$\begin{aligned} v(d(T)) &= v(d(S\cap T)) + \frac{v(d(T)) - v(d(S\cap T))}{d(T) - d(S\cap T)}[d(T) - d(S\cap T)] \\ &\geq v(d(S\cap T)) + \frac{v(d(S\cup T)) - v(d(S\cap T))}{d(S\cup T) - d(S\cap T)}[d(T) - d(S\cap T)]. \end{aligned}$$

Adding these together, we get:

$$\begin{aligned} v(d(S)) + v(d(T)) &\geq 2v(d(S\cap T)) + \frac{v(d(S\cup T)) - v(d(S\cap T))}{d(S\cup T) - d(S\cap T)} \\ &\quad \cdot [d(S) + d(T) - 2d(S\cap T)] \\ &= v(d(S\cap T)) + v(d(S\cup T)), \end{aligned}$$

as desired. □

Now we return to the pricing problem.

Theorem B.2. *Let $\bar{\pi}$ be a dual solution to $\overline{\mathcal{M}_{\mathcal{R}'}}$. For each $j \in I$, define set function g_j on $E_j = I\setminus\{j\}$ as follows. For each $S \subseteq E_j$,*

$$g_j(S) = f_j + b(S) + w_j(\mu(S)) + \theta z_\alpha \sqrt{s(S)},$$

where vectors b, μ, $s \in \Re^{E_j}$ are defined as follows:

$$b_i = \beta\mu_i d_{ij} - \bar{\pi}_i,$$
$$\mu_i = \mu_i,$$
$$s_i = \sigma_i^2,$$

for each $i \in E_j$. Then g_j is a submodular function on E_j, and the pricing problem $\mathcal{P}_j$ is exactly the problem of minimizing submodular function g_j over E_j.

Proof. By Properties B.3 and B.4, it suffices to show that each of the second, third, and fourth terms in the definition of g_j are submodular functions. The second term is submodular, by Property B.1. The third and fourth terms are submodular by Lemma B.1, together with Property B.2. □

References

Axsater, S. 1996. Using the deterministic EOQ formula in stochastic inventory control. *Management Sci.* **42** 830–834.

Barnhart, C., E. L. Johnson, G. L. Nemhauser, M. W. P. Savelsbergh, P. H. Vance. 1994. Branch-and-price: Column generation for solving huge integer programs. J. R. Birge, K. G. Murty, eds. *Mathematical Programming: State of the Art 1994*. Braun-Brumfield, Inc., Ann Arbor, MI.

Ceria, S., P. Nobili, A. Sassano. 1997. Set covering problem. M. Dell'Amico, F. Maffioli, S. Martello, eds. *Annotated Bibliographies in Combinatorial Optimization*. John Wiley and Sons, New York.

Daskin, M. S. 1995. *Network and Discrete Location: Models, Algorithms, and Applications*. Wiley-Interscience, New York.

——, S. H. Owen. 1999. Location models in transportation. R. Hall, ed. *Handbook of Transportation Science*. Kluwer Academic Publishers, Norwell, MA, 311–360.

Drezner, Z. 1995. *Facility Location: A Survey of Applications and Methods*. Springer, New York.

Edmonds, J. 1970. Submodular functions, matroids, and certain polyhedra. R. Guy, H. Hanani, N. Sauer, J. Schönheim, eds. *Combinatorial Structures and Their Applications*. Gordon and Breach, New York.

Eppen, G. 1979. Effects of centralization on expected costs in a multi-location newsboy problem. *Management Sci.* **25** 498–501.

Fergusen, T. S. 1967. *Mathematical Statistics: A Decision Theoretic Approach*. Academic Press, New York.

Received: July 2000; revision received: September 2001; accepted: July 2002.

Fujishige, S. 1991. *Submodular Functions and Optimization*. North-Holland, Amsterdam, The Netherlands.

Grötschel, M., L. Lovasz, A. Schrijver. 1981. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica* **1** 169–197.

Herer, Y., A. Rashit. 1999. Policies in a general two-location infinite horizon inventory system with lateral stock transshipments. Working paper, Israel Institute of Technology, Haifa, Israel.

——, M. Tzur, E. Yücesan. 1999. Transshipments: An emerging inventory recourse to achieve supply chain leagility. Working paper, Tel Aviv University, Israel.

Hopp, W., M. L. Spearman. 1996. *Factory Physics: Foundations of Manufacturing Management*. Irwin, Chicago, IL.

Iwata, S., L. Fleischer, S. Fujishige. 2000. A combinatorial, strongly polynomial-time algorithm for minimizing submodular functions. *Proc. 32nd Annual ACM Sympos. Theory of Comput.*, Portland, OR, 97–106.

Krishnan, K. S., V. R. K. Rao. 1965. Inventory control in *N* warehouses. *J. Indust. Engrg.* **16** 212–215.

Lovasz, L. 1982. Submodular functions and convexity. A. Bachem, M. Grötschel, B. Korte, eds. *Mathematical Programming—The State of the Art*. Springer-Verlag, New York, 235-257.

Luenberger, D. G. 1998. *Investment Science*. Oxford Unversity Press, New York.

Nahmias, S. 1997. *Production and Operations Management*. Irwin, Chicago, IL.

Nemhauser, G. L., L. A. Wolsey. 1988. *Integer and Combinatorial Optimization.* John Wiley & Sons, New York.

Owen, S. H., M. S. Daskin. 1998. Strategic facility location: A review. *Eur. J. Oper. Res.* **111** 423–447.

Rado, R. 1942. A theorem on independence relations. *Quart. J. Math.* **2** 83–89.

Schrijver, A. 1999. A combinatorial algorithm minimizing submodular functions in strongly polynomial time.

Zheng, Y. S. 1992. On properties of stochastic inventory systems. *Management Sci.* **38** 87–103.

Zipkin, P. H. 1997. *Foundations of Inventory Management*. McGraw-Hill, Boston, MA.